
W P S R L

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPSrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Fri Jun 16 08:49:52 2000; MasPar time 24.34 Seconds
Tabular output not generated. 672.401 Million cell updates/sec

Title: >US-09-360-125-1
Description: (1-347) from US09360125.pep
Perfect Score: 2538
Sequence: 1 MAQSRDGNPFAPSELDNP.....VRTAANAAGAAENAFRAP 347

Scoring table: PAM 150
Gap 11

Searched: 142080 seqs, 47172406 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: p1r62
1:p1r1 2:p1r2 3:p1r3 4:p1r4

Statistics: Mean 49.533; Variance 128.029; scale 0.387

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length DB ID	Description Pred. No.
1	2490	98.1	secretory carrier mem 0.00e+00
2	1299	51.2	secretory carrier mem 9.04e-181
3	748	29.5	hypothetical protein 1.57e-92
4	209	8.2	hypothetical protein 6.05e-12
5	141	5.6	hypothetical protein 1.46e-03
6	129	5.1	brain-specific angiot 3.15e-02
7	124	4.9	transforming protein 1.09e-01
8	125	4.9	conserved hypothetical 8.49e-02
9	116	4.6	oligopeptide transpor 7.42e-01
10	116	4.6	oligopeptide transpor 7.42e-01
11	118	4.6	ubiquinol--cytochrome 4.62e-01
12	116	4.6	NADH dehydrogenase (u 7.42e-01
13	116	4.6	NADH dehydrogenase su 7.42e-01
14	116	4.6	extensin-like protein 7.42e-01
15	117	4.6	epidermal growth fact 5.86e-01
16	118	4.6	nitrate assimilation 4.62e-01
17	115	4.5	cytochrome-c oxidase 9.39e-01
18	113	4.5	lic-1 protein B - Hae 1.50e-00
19	113	4.5	lic-1 protein B - Hae 1.50e-00
20	115	4.5	NADH dehydrogenase su 9.39e-01
21	113	4.5	preprotein translocas 1.50e-00
22	113	4.5	NADH dehydrogenase (u 1.50e-00
23	113	4.5	hydroxyproline-rich g 1.50e+00

24	115	4.5	760	2	T06291	extensin homolog T9B8	9.39e-01
25	114	4.5	948	2	A57640	retinoblastoma protei	1.19e+00
26	111	4.4	298	2	T13684	NADH dehydrogenase su	2.37e+00
27	112	4.4	373	2	S52766	secf protein - Strept	1.88e+00
28	112	4.4	475	2	A70320	nitrate transporter -	1.88e+00
29	111	4.4	496	2	T00691	hypothetical protein	2.37e+00
30	109	4.3	184	2	S78091	endocuticular protein	3.74e+00
31	108	4.3	232	2	A45612	H+-transporting ATP s	4.69e+00
32	110	4.3	265	1	E69254	cobalamin (5'-phospha	2.98e+00
33	110	4.3	299	2	A57652	cyclic nucleotide-gat	2.98e+00
34	109	4.3	421	2	S29599	acrosin (EC 3.4.21.1)	3.74e+00
35	110	4.3	439	2	C22845	NADH dehydrogenase (u	2.98e+00
36	110	4.3	448	2	T06076	extensin-like protein	2.98e+00
37	108	4.3	481	2	S47091	cyclase-associated pr	4.69e+00
38	110	4.3	530	1	G64918	phosphotransferase sy	2.98e+00
39	108	4.3	530	2	A45690	transactivator EBNA-2	4.69e+00
40	110	4.3	543	2	S25128	61K protein - Autogra	2.98e+00
41	110	4.3	543	2	H40781	hypothetical 60.7K pr	2.98e+00
42	109	4.3	705	2	S55420	conserved hypothetical	3.74e+00
43	109	4.3	804	2	S61395	probable Na+/H+-excha	3.74e+00
44	110	4.3	2715	2	T13049	evelid - fruit fly (D	2.98e+00
45	109	4.3	3951	1	VF1HB1	F1 protein - avian in	3.74e+00

ALIGNMENTS

RESULT 1
ENTRY T08826 #type complete
TITLE Secretory carrier membrane protein homolog propin1 - human
ORGANISM Homo sapiens #common_name man
DATE 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

ACCESSIONS T08826
REFERENCE 216482
#authors Winfield, S.L.; Tayebi, N.; Martin, B.M.; Ginns, E.I.; Sidransky, E.
#journal Genome Res. (1997) 7:1020-1026
#title Identification of three additional genes contiguous to the glucocerebrosidase locus on chromosome 1q21: Implications for Gaucher Disease.

#cross-references MUID:97474796
#accession T08826
#status translated from GB/EMBL/DBJ
#molecule_type DNA
#residues 1-347 #label WIN
#cross-references EMBL:AF023268; NID:g2564910; PID:g2564915

GENETICS
#gene propin1
#map_position 1
#introns 22/3; 48/3; 89/3; 130/1; 173/1; 236/2; 260/2; 299/3
#length 347 #molecular-weight 38404 #checksum 7224

SUMMARY
Query Match 98.1%; Score 2490; DB 2; Length 347;
Best Local Similarity 98.6%; Pred. No. 0.00e+00;
Matches 342; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Db 1 MAQSRDGNPFAPSELDNPQDPVAVIOHRSROYATRDVYNFETREPPAYEPAPAP 60
QY 1 MAQSRDGNPFAPSELDNPQDPVAVIOHRSROYATRDVYNFETREPPAYEPAPAP 60
Db 61 LPPPSAPLSQPSRKLSPTPEKNGYSYQASAAAATALLKKQELNKAELDREREL 120
QY 61 LPPPSAPLSQPSRKLSPTPEKNGYSYQASAAAATALLKKQELNKAELDREREL 120
Db 121 QHAALGTATQNNWPLPFCFQDPQFQDISMEIQFOKTVSTMYLWMCSTLALL 180
QY 121 QHAALGTATQNNWPLPFCFQDPQFQDISMEIQFOKTVSTMYLWMCSTLALL 180
Db 181 NFLACLASFCVETNNGAGFGLSILWLLFTPCSFVCVYRPMYKAFRSDSSNFFAFFNF 240
QY 181 NFLACLASFCVETNNGAGFGLSILWLLFTPCSFVCVYRPMYKAFRSDSSNFFAFFNF 240
Db 241 FDQDVLFLQAIIGIPGNGFSGWISALVVPKGNATVSVLMLLVALLETGTGIAVLGIVMLKRI 300

```

241 FVQDVLFLQVLAIGPFGWGSWISALVPGKNTAVSVLMLLVALLEGTAVLGIVMLKRI 300
Db 301 HSLYRTGASFOKAQOEFAAGVFNPAVTRTAANAAGAAENAFRAP 347
QY 301 HSLYRTGASFOKAQOEFAAGVFNPAVTRTAANAAGAAENAFRAP 347

RESULT 2
ENTRY #type complete
TITLE secretory carrier membrane protein 37 - rat
ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
DATE 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change
ACCESSIONS S37395; S37782
REFERENCE #molecule_type mRNA
AUTHORS Brand, S.H.; Castle, J.D.
JOURNAL EMBO J. (1993) 12:3753-3761
TITLE SCAMP 37, a new marker within the general cell surface
recycling system.
#cross-references MUID:94008982
#accession S37395
#status not compared with conceptual translation
#molecule_type mRNA
#residues 1-338 #label BRA
#accession S37782
#molecule_type protein
#residues 53-78;105-116;299-322 #label BR2
KEYWORDS calcium binding; leucine zipper; transmembrane protein; zinc
finger
FEATURE
151-176 #domain transmembrane #status predicted #label TM1
222-239 #domain transmembrane #status predicted #label TM2
258-282 #domain transmembrane #status predicted #label TM3
SUMMARY #length 338 #molecular_weight 37998 #checksum 9593

Query Match 51.2%; Score 1299; DB 2; Length 338;
Best Local Similarity 55.3%; Pred. No. 9,04e-181;
Matches 187; Conservative 71; Mismatches 63; Indels 17; Gaps 14;

Db 6 SNPFADP-DLNNPFKDSVTVTRNVP-PGLDEYNPF-S-DSR-T-PP-PGVKMPNVP 57
QY 8 GNPAEPSELONPFQDPAVQHRPSRQATLDVYAPFETREPPAYEPAPAPLPSPAP 67
Db 58 NTQPAI-MKPE--EHPAYTQITKEHALAQAEELKRQEELERKAAELDREREMON--LS 112
QY 68 SLOPSRKLSPTEPKNYSGSYSTQASAAA-ATAELLKKQELNKAELDRERELQHAALG 126
Db 113 QHG-RKNNWPPLPSNF-PVGCFCFYODFSVDIPVEFQKTVKLYLWMPHAYTLFLNIGC 170
QY 127 GTATRONNWPPLPS-FCPVPQCFODISMEIPOEFQKTVSTMYLWMCSTLALLNFLAC 185
Db 171 LAWFCVDSRAVDFGLSLILWFLLELLPCSFVCVRYLYGAFRSDSFRFVFFVYICOPA 230
QY 186 LASFCVEINNGAGFGLSLILWLLFPFCFVCVRYLYGAFRSDSFRFVFFVYICOPA 245
Db 231 VHVLAQAFHNGNCGWISSITGLNKNIPVGMIIIAALFTASAVISLWFKVHGLYR 290
QY 246 LFVLAQAFHNGNCGWISSITGLNKNIPVGMIIIAALFTASAVISLWFKVHGLYR 305
Db 291 TTGASFEKAQOEFAAGVFNPAVTRTAANAAGAAENAFRAP 328
QY 306 RTGASFEKAQOEFAAGVFNPAVTRTAANAAGAAENAFRAP 343

RESULT 3
ENTRY #type fragment
TITLE hypothetical protein M01D7.2 - Caenorhabditis elegans
ORGANISM #formal_name Caenorhabditis elegans
DATE 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change
ACCESSIONS T15283
```

```

REFERENCE 218322
AUTHORS Gattung, S.; Goela, D.; Wilson, R.
SUBMISSION submitted to the EMBL Data Library, May 1997
DESCRIPTION The sequence of C. elegans cosmid M01D7.
ACCESSION T15283
STATUS preliminary; translated from GB/EMBL/DDBJ
#molecule_type DNA
#residues 1-324 #label GAT
#cross-references EMBL:AF003739; NID:g2105482; PID:g2105487;
#experimental_source strain Bristol N2; clone M01D7
GENETICS
#gene CESP:M01D7.2
#map_position 1
SUMMARY #length 324 #checksum 8640

Query Match 29.5%; Score 748; DB 2; Length 324;
Best Local Similarity 43.1%; Pred. No. 1.57e-92;
Matches 110; Conservative 70; Mismatches 63; Indels 12; Gaps 10;

Db 66 DLERRAQLRMREELDRRQSAAGNNLNTNAQNNAPRPHNPPLPTIPIEPCFYQDI 125
QY 98 ELLKQELNKAELDRRERELQHAALG-CT---ATRONNWPPLPSCFVQPCFFQDI 152
Db 126 EVEIPVQKTVTFAYVYVLMYVLAIVNVLASL-FYMFAGSGIGQLFLACIQALFSP 184
QY 153 SMEIPOEFQKTVSTMYLWMCSTLALLNFLACLSFCVETNNGAG-FGLSILWLLFTP 211
Db 195 CSFLFWFPVYKAFRNDSENFVFFVFFVFFVFFVFFVFFVFFVFFVFFVFFVFFV 242
QY 212 CSFVCWYRPMYKAFRSDSFRFVFFVFFVFFVFFVFFVFFVFFVFFVFFVFFV 270
Db 243 VSVPTALMLLSAIFRT-VALTGMVTVLVKVRHLYRGAGFSDKARQETNGVMSDAGVQ 301
QY 271 GNTAVSVLMLLVALLEGTAVLGIVM-LKRIHSLYRRTGASFOKAQOEFAAGVFNPAVR 329
Db 302 RATQAATQAAGAAF 316
QY 330 TAAANAAGAAENAF 344

RESULT 4
ENTRY #type complete
TITLE hypothetical protein F21B7.17 - Arabidopsis thaliana
ORGANISM #formal_name Arabidopsis thaliana #common_name mouse-ear
cress
DATE 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change
ACCESSIONS T00903
REFERENCE 214208
AUTHORS Shinn, P.; Buehler, E.; Dewar, K.; Feng, J.; Kim, C.; Li, Y.;
Sun, H.; Conway, A.; Conway, A.; Kurtz, D.; Ojil, O.; Shen,
Y.K.; Toriumi, M.; Vysotskaia, V.; Yu, G.; Davis, R.W.;
Federapfel, N.A.; Theologis, A.; Ecker, J.R.
SUBMISSION submitted to the EMBL Data Library, January 1998
DESCRIPTION Genomic sequence for Arabidopsis thaliana BAC F21B7.
ACCESSION T00903
STATUS preliminary; translated from GB/EMBL/DDBJ
#molecule_type DNA
#residues 1-221 #label SHI
#cross-references EMBL:AC002560; NID:g2618677; PID:g2809248;
GSPDB:GN0059; ATSP:F21B7.17
GENETICS
#gene ATSP:F21B7.17
#map_position 1
#intons 18/3; 53/3; 74/3; 86/1; 117/2; 132/3; 155/2; 171/3; 193/3
SUMMARY #length 221 #molecular_weight 24765 #checksum 5742

Query Match 8.2%; Score 209; DB 2; Length 221;
Best Local Similarity 30.4%; Pred. No. 6.05e-12;
Matches 34; Conservative 38; Mismatches 36; Indels 4; Gaps 2;

Db 92 LSIIYFLAGVPGAVVLTWRPLRYRATRTDSALKFGAFFFFVFFVFFVFFVFFVFFV 151
```

Result No.	Query		DB	ID	Description	Pred. No.
	Score	Match Length				
1	1397	57.4	338	S37395	secretory carrier mem	5.93e-198
2	1223	50.2	347	T08826	secretory carrier mem	1.30e-169
3	670	27.5	334	T15883	hypothetical protein	7.18e-81
4	246	10.1	221	T00903	hypothetical protein	4.56e-17
5	128	5.3	684	T13491	NADH dehydrogenase	3.45e-02
6	129	5.3	731	T12226	NADH dehydrogenase	2.69e-02
7	126	5.2	700	T13471	NADH dehydrogenase su	5.71e-02
8	124	5.1	686	T12125	NADH dehydrogenase	9.37e-02
9	124	5.1	658	T12560	NADH dehydrogenase su	9.37e-02
10	122	5.0	285	S73885	MG135 homolog A65_orf	1.53e-01
11	121	5.0	288	T13684	NADH dehydrogenase su	1.96e-01
12	122	5.0	684	T12147	NADH dehydrogenase	1.53e-01
13	121	5.0	698	T12586	NADH dehydrogenase su	1.96e-01
14	119	4.9	638	1 QOB2Y21	mRNA maturase b14 - y	3.18e-01
15	119	4.9	736	2 T12114	NADH dehydrogenase	3.18e-01
16	119	4.9	740	2 T12223	NADH dehydrogenase	3.18e-01
17	118	4.8	383	2 C71683	rod shape-determining	4.04e-01
18	116	4.8	398	1 A229331	ubiquinol--cytochrome	6.51e-01
19	118	4.8	648	2 T12570	NADH dehydrogenase su	4.04e-01
20	116	4.8	683	2 T12149	NADH dehydrogenase	6.51e-01
21	116	4.8	683	2 T12149	NADH dehydrogenase	6.51e-01
22	117	4.8	686	2 T12128	NADH dehydrogenase	5.13e-01
23	117	4.8	687	2 T12126	NADH dehydrogenase	5.13e-01

QY 178 SSKGVDFGLSILWELIETPCAEFCWYRPIYKAERSDNSESEFVFEFFCOIGIYTIQIV 237

```

Db 238 GFHNKNGCGWISSLTGLNKN-IPVGMIMIIITAAETASAVISLVNFKVKGHLRYRTGASG 296
Qy 238 GIPGLGDSGWAALSTLDNLSAIVSMVWVAGFTLCVLSVFLQRVHSLYRRTGASG 297
Db 297 EKAQGEFATGVMSNKTIVTAAANAATA 324
Qy 298 QAOQEESQIGFSRTFTHRAASSAQA 325

RESULT 2
ENTRY #type complete
TITLE secretory carrier membrane protein homolog propin1 - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change
20-Sep-1999

ACCESSIONS T08826
REFERENCE Winfield, S.L.; Tayebi, N.; Martin, B.M.; Ginns, E.I.;
#authors Stofransky, E.
#journal Genome Res. (1997) 7:1020-1026
#title Identification of three additional genes contiguous to the
glucocorticoidase locus on chromosome 1q21: Implications
for Gaucher Disease.
#cross-references MUID:97474796
#accession T08826
#status translated from GB/EMBL/DBJ
#molecule_type DNA
#residues 1-347 #label WTN
#cross-references EMBL:AF032268; NID:g2564910; PID:g2564915
GENETICS propin1
#gene
#map_position 1
#introns 22/3; 48/3; 130/1; 173/1; 226/2; 260/2; 299/3
SUMMARY #length 347 #molecular_weight 38404 #checksum 7224

Query Match 50.2%; Score 1223; DB 2; Length 347;
Best Local Similarity 52.6%; Pred. No. 1.36e-169;
Matches 175; Conservative 72; Mismatches 70; Indels 16; Gaps 12;

Db 9 NPFAPESELDNPFDDPAVQHRPQRYATRDVNPFFETREPPAYEPAPAPLPPPSAPS 68
Qy 7 NPFAADPVV-NPFDPSVTQLTNAPQGLAE-FNPF-S-ETNAATTVPV-TQLPGSSQA 61
Db 69 -LQPSRLSPTEPNKYSYSQASAAATAELLKKQELNKAELDRRELQHALGG 127
Qy 62 VLQPS--VEPTQ-----T-P-QAVVSAQAQLLQROELDRKAAELERKRELQNT-VAN 112
Db 128 TATRONNPPLPSPQVPCPFODISMEIPOEFQKTVSTMYLWMCSTLALLNPLACLA 187
Qy 113 LHVRRNNWPPLPSCWPKPCFYQDFSTEIPADYORICKMXYLWMLHSVTILFLNLLACLA 172
Db 188 SFCVETNNGAGGSLWLLFTPCSFVGVWRPMYKAFRSDSSNFNFAFFNFDDQDLF 247
Qy 173 WFGNSKSGVDGSLIWLFTPCFLCFLCWRYPIYKAFRSDNSFFVFFVFCQIGY 232
Db 248 VLQAGTGPWCFSGWISALVVPKNT-AVSVMLLVALLFTGIAVLGVLWMLKRHSLYR 306
Qy 233 IQLVGLPGLGDSGWIAALSTLDNLSAIVSMVWVAGFTLCVLSVFLQRVHSLYR 292
Db 307 TGASFQQAQGEFAAGVFSNPVTRRAANAAGA 339
Qy 293 TGASFQQAQGEFSQIGFSRTFTHRAASSAQA 325

RESULT 3
ENTRY #type fragment
TITLE T15283
hypoetical protein M01D7.2 - Caenorhabditis elegans
ORGANISM #formal_name Caenorhabditis elegans
DATE 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change
20-Sep-1999
T15283
ACCESSIONS

```

```

REFERENCE 218222
#authors Gattung, S.; Goela, D.; Wilson, R.
#submission submitted to the EMBL Data Library, May 1997
#description The sequence of C. elegans cosmid M01D7.
#accession T15283
#status Preliminary; translated from GB/EMBL/DBJ
#molecule_type DNA
#residues 1-324 #label GAT
#cross-references EMBL:AF003739; NID:g2105482; PID:g2105487;
PIDN:AAB58069.1; GSPDB:GN00019; GESP:M01D7.2
#experimental_source strain Bristol N2; clone M01D7
GENETICS
#gene CESP:M01D7.2
#map_position 1
SUMMARY #length 324 #checksum 8640

Query Match 27.5%; Score 670; DB 2; Length 324;
Best Local Similarity 36.9%; Pred. No. 7.18e-81;
Matches 109; Conservative 71; Mismatches 102; Indels 13; Gaps 12;

Db 25 DFNPFANRAGSNQPTTHQSTGSLGNKSGAGMDELFRKQODLERRAQE-LNMREEELDR 83
Qy 36 EFNPFSE-TNAATTVPV-TQLPGSSQPAVLQPSVEPTQTPQAVVSAQAAGLLRQOEELDR 94
Db 84 RORSAA-GGNLNTNQNNAAPRPHNPPLPIIPIEPCFYQDIEVEIPVQKQTFYFY 142
Qy 95 KAELEKEREKELQNTVANLHVRRNNWPPLPSCWPKPCFYQDFSTEIPADYORICKMLY 154
Db 143 VFLMYVALVNVNLSLAFYMFAG-GSIGOLF-LACIOLALFSPCSFELFWPVPVKAERN 200
Qy 155 LWMLHSVTILFLNLLACLAW--FSGNSKGVDFGSLWFLTPCFLCWRYPIYKAFRS 212
Db 201 DSSNFVFFVFLFPHC-IFTFV-QTLGLSNTA-CGWINALDTF-NVSVPIALLMLISAI 256
Qy 213 DNSFFVFFVFFV-FF-CQIGIVIIQLVGPGLGDSGWIAALSTLDNLSAIVSMVWVAG 270
Db 257 TFTVALGWTALVKVHLYRGAGFSDIKARQETNGVNSDAGVORATQAATQAA 311
Qy 271 FFTLCVLSVFLQRVHSLYRRTGASEFQQAQGEFSQIGFSRTFTHRAASSAQA 325

RESULT 4
ENTRY #type complete
TITLE hypothetical protein F21B7.17 - Arabidopsis thaliana
ORGANISM #formal_name Arabidopsis thaliana #common_name mouse-ear
cress
DATE 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change
20-Sep-1999
ACCESSIONS T00903
REFERENCE Z14208
#authors Shinn, P.; Buehler, E.; Dewar, K.; Feng, J.; Kim, C.; Li, Y.;
Sun, H.; Conway, A.; Conway, A.; Kurtz, D.; Oji, O.; Shen,
Y.K.; Toriumi, M.; Vysotskaia, V.; Yu, G.; Davis, R.W.;
Fedorpiel, N.A.; Theologis, A.; Ecker, J.R.
#submission submitted to the EMBL Data Library, January 1998
#description Genomic sequence for Arabidopsis thaliana BAC F21B7.
#accession T00903
#status Preliminary; translated from GB/EMBL/DBJ
#molecule_type DNA
#residues 1-221 #label SHI
#cross-references EMBL:AC002560; NID:g2618677; PID:g2809248;
GSPDB:GN0059; ATSP:F21B7.17
GENETICS
#gene ATSP:F21B7.17
#map_position 1
#introns 18/3; 53/3; 74/3; 86/1; 117/2; 132/3; 155/2; 171/3; 193/3
SUMMARY #length 221 #molecular_weight 24765 #checksum 5742

Query Match 10.1%; Score 246; DB 2; Length 221;
Best Local Similarity 33.3%; Pred. No. 4.56e-17;
Matches 38; Conservative 37; Mismatches 32; Indels 7; Gaps 4;

Db 92 LSIIFYLAGVPGAVLVLRPLRYRATRTDSALKFGAFFFYVFIHAFCGFAA-VAPPVIFQ 150

```

M P S R L H
***** (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd.
MPsrch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Fri Jun 16 08:48:16 2000; MasPar time 15.54 Seconds
Tabular output not generated. 680.011 Million cell updates/sec

Title: >US-09-360-125-1
Description: (1-347) from US09360125.pep
Perfect Score: 2538
Sequence: 1 MAQSRDGGNPAEPSELNDP.....VNTAANAAGAAENAFRAP 347

Scoring table:
PAM 150
Gap 11

Searched: 83857 seqs, 30454973 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: swiss-prot38
1:swissprot

Statistics: Mean 50.311; Variance 114.866; scale 0.438

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match	Length	ID	Description	Pred. No.
1	2490	98.1	347	1	SCA3_HUMAN SECRETORY CARRIER-ASSO	0.00e+00
2	2321	91.4	349	1	SCA3_MOUSE SECRETORY CARRIER-ASSO	0.00e+00
3	1299	51.2	338	1	SCA1_RAT SECRETORY CARRIER-ASSO	1.11e-204
4	1279	50.4	338	1	SCA1_HUMAN SECRETORY CARRIER-ASSO	5.11e-201
5	1256	49.5	328	1	SCA2_HUMAN SECRETORY CARRIER-ASSO	8.29e-197
6	124	4.9	323	1	JUND_CHICK TRANSCRIPTION FACTOR J	1.91e-02
7	125	4.9	410	1	HID_DROME HEAD INVOLUTION DEFECT	1.45e-02
8	119	4.7	579	1	YN05_CAEEL PUTATIVE SODIUM- AND C	7.35e-02
9	116	4.6	302	1	OPPC_ECOLI OLIGOPEPTIDE TRANSPORT	1.62e-01
10	116	4.6	302	1	OPPC_SALTY OLIGOPEPTIDE TRANSPORT	1.62e-01
11	117	4.6	822	1	EPS8_HUMAN EPIDERMAL GROWTH FACTO	1.25e-01
12	118	4.6	892	1	MIRA_EMENT NITROGEN ASSIMILATION	9.59e-02
13	113	4.5	292	1	LICB_HAEIN LICB PROTEIN.	3.54e-01
14	113	4.5	620	1	EXTN_TOBAC EXTENSIN PRECURSOR (CE	3.54e-01
15	114	4.5	3110	1	HD_RAT HUNTINGTIN (HUNTINGTON	2.74e-01
16	112	4.4	373	1	SECF_STRCO PROTEIN-EXPORT MEMBRAN	4.58e-01
17	111	4.4	496	1	YML5_ARATH HYPOTHETICAL MLO-LIKE	5.91e-01
18	111	4.4	527	1	BBF1_CANAL TRANSCRIPTION FACTOR R	5.91e-01
19	112	4.4	899	1	PMAL_KLULA PLASMA MEMBRANE ATPASE	4.58e-01
20	108	4.3	175	1	Y433_METJA HYPOTHETICAL PROTEIN M	1.26e+00
21	108	4.3	305	1	NULM_APILI NADH-UBIQUINONE OXIDOR	1.26e+00
22	110	4.3	464	1	GUALEIDO GUANINE NUCLEOTIDE-BIN	7.62e-01
23	108	4.3	481	1	CAP_CHLVR ADENYLYL CYCLASE-ASSOC	1.26e+00

RESULT 1
ID SCA3_HUMAN STANDARD; PRT; 347 AA.
AC O14828; O15128;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE SECRETORY CARRIER-ASSOCIATED MEMBRANE PROTEIN 3.
GN SCAMP3 OR PROPIN1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP TISSUE FROM N.A.
RC TISSUE-BRAIN;
RX MEDLINE; 97474796.
RA Winfield S.L., Tayebi N., Martin B.M., Ginns E.I., Sidransky E.;
RT Identification of three additional genes contiguous to the
RT glucocerebrosidase locus on chromosome 1q21: implications for Gaucher
RT disease.";
RL Genome Res. 7:1020-1026(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 98017831.
RA Singleton D.R., Wu T.T., Castle J.D.;
RT Three mammalian SCAMPs (secretory carrier membrane proteins) are
RT highly related products of distinct genes having similar subcellular
RT distributions.";
RL J. Cell Sci. 110:2099-2107(1997).
CC -1- FUNCTION: FUNCTIONS IN POST-GOLGI RECYCLING PATHWAYS. ACTS AS A
CC RECYCLING CARRIER TO THE CELL SURFACE.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED WITH HIGHEST EXPRESSION IN
CC HEART AND SKELETAL MUSCLE.
CC -1- SIMILARITY: BELONGS TO THE SCAMP FAMILY.
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to licensed@isb-sib.ch).
CC -----
CC ENBL; AF023268; AAC51821.1; -
CC ENBL; AF005039; AAB62724.1; -
DR Transmembrane; Transport; Protein transport; Multigene family.
KW TRANSMEM 171 191
FT POTENTIAL.

ALIGNMENTS

24	108	4.3	485	1	SSGP_VOLCA	SULFATED SURFACE GLYCO	1.26e+00
25	109	4.3	491	1	Y039_HUMAN	HYPOTHETICAL PROTEIN K	9.81e-01
26	108	4.3	497	1	NU4M_ACACA	NADH-UBIQUINONE OXIDOR	1.26e+00
27	110	4.3	530	1	PTOA_ECOLI	PTS SYSTEM, MALTOSE AN	7.62e-01
28	110	4.3	543	1	VP61_NPVAC	61 KD PROTEIN.	7.62e-01
29	109	4.3	705	1	YWJF_BACSU	HYPOTHETICAL 79.2 KD P	9.81e-01
30	108	4.3	815	1	MK07_HUMAN	MITOGEN-ACTIVATED PROT	1.26e+00
31	108	4.3	947	1	PMA2_YEAST	PLASMA MEMBRANE ATPASE	1.26e+00
32	109	4.3	3951	1	VGFI_IBVB	F1 PROTEIN.	9.81e-01
33	106	4.2	125	1	SECG_BORBU	PROBABLE PROTEIN-EXPO	2.07e+00
34	106	4.2	257	1	YXP2_XANCP	HYPOTHETICAL 26.9 KD P	2.07e+00
35	107	4.2	389	1	NDPP_MOUSE	NPC DERIVED PROLINE RI	1.62e+00
36	106	4.2	389	1	FL_ORISA	PUTATIVE TRANSCRIPTION	2.07e+00
37	107	4.2	412	1	ALF_PETHY	ALF PROTEIN (ABERRANT	1.62e+00
38	107	4.2	426	1	EXLP_TOBAC	PISTIL-SPECIFIC EXTEN	1.62e+00
39	107	4.2	443	1	HXA3_HUMAN	HOMEOBOX PROTEIN HOX-A	1.62e+00
40	106	4.2	444	1	NU4M_LOCM1	NADH-UBIQUINONE OXIDOR	2.07e+00
41	106	4.2	473	1	NORB_PSEST	NITRIC-OXIDE REDUCTASE	2.07e+00
42	107	4.2	551	1	YD25_YEAST	HYPOTHETICAL 62.6 KD P	1.62e+00
43	106	4.2	635	1	VP40_HSV11	CAPSID PROTEIN P40 (VI	1.62e+00
44	107	4.2	1744	1	TENS_CHICK	TENSIN.	1.62e+00
45	105	4.1	364	1	FTSW_BORBU	CELL DIVISION PROTEIN	2.65e+00

FT TRANSMEM 197 217 POTENTIAL.
FT TRANSMEM 247 267 POTENTIAL.
FT TRANSMEM 277 297 POTENTIAL.
FT CONFLICT 3 3 O -> R (IN REF. 2).
FT CONFLICT 38 38 R -> L (IN REF. 2).
FT CONFLICT 74 74 K -> M (IN REF. 2).
FT CONFLICT 235 235 A -> V (IN REF. 2).
FT CONFLICT 239 239 N -> I (IN REF. 2).
FT CONFLICT 242 242 D -> V (IN REF. 2).
FT CONFLICT 331 331 R -> A (IN REF. 2).
SQ SEQUENCE 347 AA; EC55F3793D20E58C CRC64;

Query Match 98.18; Score 2490; DB 1; Length 347;
Best Local Similarity 98.64; Pred. No. 0.00e+00;
Matches 342; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Db 1 MAQRDGNFPAESELNDNFQDPAVQHRSROYATDVNPNPTEPREPPAYEPAPAP 60
QY 1 MAQRDGNFPAESELNDNFQDPAVQHRSROYATDVNPNPTEPREPPAYEPAPAP 60

Db 61 LPPPSAPSLQSRKLSPTPEKNTGYSYQASAAAATAELLKKQELNRKAEELDRREL 120
QY 61 LPPPSAPSLQSRKLSPTPEKNTGYSYQASAAAATAELLKKQELNRKAEELDRREL 120

Db 121 QHAALGATATRONNWPPLPSPFCVQPCFQDIMEIPQEFQKTVSTMYLWMCSTLALL 180
QY 121 QHAALGATATRONNWPPLPSPFCVQPCFQDIMEIPQEFQKTVSTMYLWMCSTLALL 180

Db 181 NFLACLASFCVETNNGAGFGLSILWLLFTPCSFVQWYRPMYKAFRSDSSENFVFFNF 240
QY 181 NFLACLASFCVETNNGAGFGLSILWLLFTPCSFVQWYRPMYKAFRSDSSENFVFFNF 240

Db 241 FQDVLFLVQAIQIPGFGFSGWISALVYKGTAVSVLMLLVALFTGIAVLGIVMLKRI 300
QY 241 FQDVLFLVQAIQIPGFGFSGWISALVYKGTAVSVLMLLVALFTGIAVLGIVMLKRI 300

Db 301 HSLYRTGASFOKAQOFAAGVFNPAVTRTAANAAGAAENAFAP 347
QY 301 HSLYRTGASFOKAQOFAAGVFNPAVTRTAANAAGAAENAFAP 347

RESULT 2
ID SCAL_RAT STANDARD; PRT; 349 AA.
AC Q35609;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE SECRETORY CARRIER-ASSOCIATED MEMBRANE PROTEIN 3.
GN SCAMP3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
[1]
SEQUENCE FROM N.A.
MEDLINE: 98017831.
RA Singleton D.R., Wu T.T., Castle J.D.;
RT Three mammalian SCAMPs (secretory carrier membrane proteins) are highly related products of distinct genes having similar subcellular distributions.*;
RL J. Cell Sci. 110:2099-2107(1997).
CC -1- FUNCTION: FUNCTIONS IN POST-GOLGI RECYCLING PATHWAYS. ACTS AS A RECYCLING CARRIER TO THE CELL SURFACE.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE SCAMP FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC EMBL: L22079; ; NOT ANNOTATED.CDS
DR Transmembrane; Transport; Protein transport; Multigene family.
FT TRANSMEM 156 176 POTENTIAL.
FT TRANSMEM 182 202 POTENTIAL.

DR EMBL; AF005036; AAB62721.1; --
KW Transmembrane; Transport; Protein transport; Multigene family.
FT TRANSMEM 169 189 POTENTIAL.
FT TRANSMEM 200 220 POTENTIAL.
FT TRANSMEM 236 256 POTENTIAL.
FT TRANSMEM 277 297 POTENTIAL.
SQ SEQUENCE 349 AA; 38387 MW; D54A514769A49309 CRC64;

Query Match 91.48; Score 2321; DB 1; Length 349;
Best Local Similarity 89.48; Pred. No. 0.00e+00;
Matches 312; Conservative 23; Mismatches 12; Indels 2; Gaps 1;

Db 1 MAQRDGTGNFSDSGLNDNFQDPAVQHRSQOYATDVNPNFENREPPAYEPAPAP 60
QY 1 MAQRDGTGNFSDSGLNDNFQDPAVQHRSQOYATDVNPNFENREPPAYEPAPAP 60

Db 61 APIPPSAPSVQSRKLSPTPEKNTGYSYQASAAAATAELLKKQELNRKAEELDRER 120
QY 61 --LPPPSAPSVQSRKLSPTPEKNTGYSYQASAAAATAELLKKQELNRKAEELDRER 118

Db 121 ELQHVAGGAGTATRONNWPPLPSPFCVQPCFQDIMEIPQEFQKTVSTMYLWMCSTLAL 180
QY 121 ELQHVAGGAGTATRONNWPPLPSPFCVQPCFQDIMEIPQEFQKTVSTMYLWMCSTLAL 178

Db 181 LLNFFACLARFCVDTGGSGFGLSMLWLLFTPCSFVQWYRPMYKAFRSDSSENFVFFNF 240
QY 179 LLNFFACLARFCVDTGGSGFGLSMLWLLFTPCSFVQWYRPMYKAFRSDSSENFVFFNF 238

Db 241 IFVQDVLFLVQAIQIPGFGFSGWISALVYKGTAVSVLMLLVALFTGIAVLGIVMLK 300
QY 239 IFVQDVLFLVQAIQIPGFGFSGWISALVYKGTAVSVLMLLVALFTGIAVLGIVMLK 298

Db 301 RIHSLYRTGASFOKAQOFAAGVFNPAVTRTAANAAGAAENAFAP 349
QY 299 RIHSLYRTGASFOKAQOFAAGVFNPAVTRTAANAAGAAENAFAP 347

RESULT 3
ID SCAL_RAT STANDARD; PRT; 338 AA.
AC P56603;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE SECRETORY CARRIER-ASSOCIATED MEMBRANE PROTEIN 1 (SCAMP 37).
GN SCAMP1 OR SCAMP.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
[1]
SEQUENCE FROM N.A.; AND PARTIAL SEQUENCE.
RP STRAIN-SPRAGUE-DAWLEY; TISSUE=BRAIN;
RX MEDLINE: 94008982.
RA Brand S.H., Castle J.D.;
RT "SCAMP 37, a new marker within the general cell surface recycling system.*";
RL EMBO J. 12:3753-3761(1993).
CC -1- FUNCTION: FUNCTIONS IN POST-GOLGI RECYCLING PATHWAYS. ACTS AS A RECYCLING CARRIER TO THE CELL SURFACE.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE SCAMP FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC EMBL: L22079; ; NOT ANNOTATED.CDS
DR Transmembrane; Transport; Protein transport; Multigene family.
FT TRANSMEM 156 176 POTENTIAL.
FT TRANSMEM 182 202 POTENTIAL.